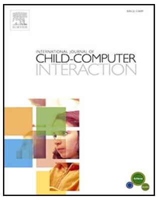
International Journal of Child-Computer Interaction 48 (2026) 100822

Contents lists available at ScienceDirect

International Journal of Child-Computer Interaction

journal homepage: www.elsevier.com/locate/ijcci



Research Paper

Preschoolers attribute preferences in response to human but not robot gaze

Federico Manzi [a,b,c,h,1,*], Mitsuhiko Ishikawa [d,1], Cinzia Di Dio [a,b], Shoji Itakura [e], Takayuki Kanda [f,g], Hiroshi Ishiguro [g,h], Davide Massaro [a,b], Antonella Marchetti [a,b]

[a] Research Center on Theory of Mind and Social Competences in the Lifespan, Department of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy
[b] Research Unit on Psychology and Robotics in the Lisfespan, Department of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy
[c] IRCCS, Fondazione Don Carlo Gnocchi, Milan, Italy
[d] Institute for Advanced Study, Hitotsubashi University, Tokyo, Japan
[e] Ritsumeikan University, Osaka, Japan
[f] Human-Robot Interaction Laboratory, Department of Computer Science, Kyoto University, Kyoto, Japan
[g] Advanced Telecommunications Research Institute International, IRC/HIL, Keihana Science City, Kyoto, Japan
[h] Department of System Innovation, Osaka University, Toyonaka, Japan

ARTICLE INFO

Keywords:
Gaze cueing
Preference formation
Attribution of preference
Theory of mind
Mental states
Child-robot interaction

ABSTRACT

With technological advancements, children increasingly interact with robots designed to mimic human-like behaviors for communication, among which gaze is particularly pivotal from early childhood. This study thus explores how children attribute and form preferences when exposed to human versus robotic gazes. The research involved 58 Italian children aged 3 to 5 years. They watched videos featuring a human and a robot each gazing at one of two objects. Subsequently, children were asked which object the gazer preferred (preference attribution) and to indicate their own preference (preference formation). Attribution of object preference was evaluated also as a function of children's Theory of Mind (i.e., false belief) and mental state attributions to human and robot agents. Results showed that children consistently attributed preferences based on human gaze, but not robot gaze, suggesting that they interpret human gaze as a meaningful communicative signal, likely associated with intentionality. Gaze had no significant effect on children's own preferences for either agent. Importantly, attribution of mental states to the human, but not to the robot, significantly predicted accurate preference attribution. No associations were found between performance on the false-belief task and gaze-based responses, indicating that explicit preference attribution may rely on socio-cognitive processes distinct from belief-based reasoning. These findings provide design-relevant insights for child–robot interaction, suggesting that gaze alone may not function as an effective communicative cue for young children and highlighting the importance of developmentally informed interaction strategies in robotic systems designed for early childhood.

1. Introduction

From the early stage of life, humans have a high sensitivity to another's gaze direction (Farroni et al., 2004). This social cues is one of the most important pieces of information for social interaction and humans engage in gaze interaction from infancy (Emery, 2000). Studies have shown that by 6 months infants follow others' gaze towards an object, suggesting that they learn the environment surrounding them through these interactions (Senju et al., 2008). In addition, EEG studies have shown that another's gaze cueing to an object facilitates information processing in infants (Hoehl et al., 2008; Striano et al., 2006). Longitudinal studies have revealed that gaze following in infancy predicts vocabulary size in later life, suggesting that following another's gaze and looking at the same object are related to word learning (Brooks and Meltzoff, 2005, 2008). Thus, processing another's gaze direction towards an object may be a core of social learning from infancy (Csibra and Gergely, 2009). Crucially, gaze following is not only an orienting response but can also reflect an emerging understanding that gaze is referential—conveying information about objects in the world. For example, Csibra and Volein (2008) demonstrated that infants can infer the presence of hidden objects based on referential gaze information, suggesting that infants use gaze cues to generate expectations about

---

* Corresponding author. Research Center on Theory of Mind and Social Competences in the Lifespan, Department of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy.

E-mail address: federico.manzi@unicatt.it (F. Manzi).

[1] These authors contributed equally.

https://doi.org/10.1016/j.ijcci.2026.100822
Received 16 July 2025; Received in revised form 27 March 2026; Accepted 14 April 2026

object presence even when the object itself is not directly visible. Moreover, individual differences in how infants allocate attention during gaze-following episodes are meaningfully related to later developmental outcomes. Parsons et al. (2019) showed that infants who later developed autism spectrum disorder followed gaze as frequently as typically developing peers, but spent less time attending to the referred objects, and that such differences in attention allocation were associated with reduced verbal abilities. Consistent with these findings, a recent systematic review concluded that eye gaze facilitates infant learning by modulating attention, arousal, and memory, and discussed whether gaze functions as a specialized socio-cognitive cue rather than merely a low-level attentional signal (Çetinçelik et al., 2021).

To examine the impact of another individual's gaze on cognitive processing, gaze cueing paradigms are commonly employed in controlled experimental settings. In these paradigms, participants observe a face that gazes toward a specific object, while other objects remain ungazed. Through repeated exposure to such scenarios, participants can learn to associate the gaze direction with the presence or absence of attention to specific objects. Gaze cueing can facilitate not only object information processing but also affect preferences for the gazed objects. Ten-month-old infants preferentially choose objects gazed at by another more often than non-gazed objects after observing gaze cueing situations (Ishikawa et al., 2019). Also, infants preferentially look at the face gazed at by another longer than the non-gazed face (Ishikawa & Itakura, 2018). These studies have suggested that another's gaze direction affects preferences for the gaze targets in infants. Similar gaze effects have been reported in adult studies. For example, Baylis et al. (2006) presented gaze cueing situations and asked how much the participants liked objects which were gazed at or non-gazed. They found that participants significantly preferred the gazed objects compared to non-gazed objects.

The mechanisms of why another's gaze direction modulates object preferences have been argued. It has been hypothesised that gaze effects on object preferences may be related to the understanding of the mental states of the social partner while she/he is gazing at the target. Gaze direction can be a signal for liking, interest or desire (Baron-Cohen et al., 1995). Therefore, during observing gaze cueing situations, humans may attribute mental states to the social partner and use gaze information for evaluations of objects (Bayliss et al., 2006). Doherty et al. (2015) investigated the development of referential gaze understanding from 2- to 4-year-olds using luminance cues, geometrical cues and normal face cues. They showed that gaze following of 2- to 3-year-olds seems to be based on the luminance information in the eye, while the older age group used both luminance and geometric information for gaze following. Doherty et al. (2015) suggested that understanding of other people as agents with mental states drives the development of the use of more precise geometric information about gaze direction. Thus, the attribution of mental states to others may be related to the effects of another's gaze direction. In another adult study, emotional expressions of cueing faces were manipulated to examine how another's mental states modulate gaze effects on object preferences (Bayliss et al., 2007). They found that objects gazed at by the happy face were liked more than objects gaze at by the disgust face, suggesting that the gaze effects on object preferences include the process of another's mental states.

Given the increasing presence of humanoid robots in social and educational settings (Belpaeme et al., 2018), it is important to understand how children interpret social cues like gaze from robots. From a Child–Computer Interaction (CCI) perspective, understanding how children interpret gaze cues from artificial agents is important not only for social interaction but also for learning outcomes (Okumura et al., 2013; O'Connell et al., 2009). This question aligns with core CCI research themes concerning how children interact with and make sense of emerging interactive technologies, including social robots and their communicative behaviors (Giannakos et al., 2020). Gaze is widely implemented in educational and social robots to guide attention, establish joint attention, and scaffold learning processes (for a review see Admoni and Scassellati, 2017). Prior work shows that even minimal head and gaze movements can shape children's perception of a robot's animacy, likeability, and helpfulness (Zaga et al., 2017). At the same time, deictic behaviors such as pointing require careful task-sensitive design, as they can sometimes hinder child–robot collaboration (Yadollahi et al., 2018). However, it remains unclear whether gaze alone constitutes a sufficient communicative signal for young children to infer intentions and preferences from robotic agents, or whether additional interactive behaviors are necessary to support effective learning. Identifying which robot behaviors are sufficient, and which require integration with richer multimodal cues, is essential for the age-sensitive design of child–robot interaction systems. The present study addresses this issue by examining how preschool children interpret gaze cues from human and robot agents in relation to mental state attribution and ToM abilities.

Developmental studies indicate that following a robot's gaze does not necessarily imply that children treat robot gaze as referential in the same way as human gaze. O'Connell et al. (2009) showed that although infants could follow the gaze of humanoid robot toward a potential referent, they failed to establish word–object associations when the gazing agent was that robot. This dissociation suggests that gaze following alone is insufficient for social learning when the agent lacks perceived psychological relevance. Complementing this finding, Meltzoff et al. (2010) demonstrated that infants' gaze following toward a humanoid robot depends on prior socially contingent interactions, suggesting that infants may treat robots as psychological agents only under specific conditions. Together, these findings highlight the importance of distinguishing between basic attentional orienting and higher-level inferences about internal states when evaluating gaze cues from robots. Complementing these developmental findings, CCI research indicates that preschool children can use robots' non-verbal cues—specifically gaze direction and body orientation—to support word learning when embedded in an interactive learning context (Kory Westlund et al., 2017). The humanoid robot Robovie, used in the present study, has been employed in free-form classroom activities to stimulate curiosity toward scientific subjects (Shiomi et al., 2015) and to promote peer relationships among elementary school children (Kanda et al., 2007). Other humanoid robots, such as NAO, have been used to support second language acquisition in primary school classrooms, including with migrant children (Konijn et al., 2022), and to foster cooperative play and social interaction (Belpaeme et al., 2018). Additionally, iCub and QT have been used in special needs education settings with children with autism spectrum disorder, supporting communicative and social behaviors (El-Muhammady et al., 2023; Ghiglino et al., 2023). More recent work in the CCI literature has further shown that humanoid robots can support early lexical learning processes in children, highlighting the potential of socially interactive robots as learning partners (Esfandbod et al., 2023). Together, these examples illustrate the growing role of humanoid robots in supporting both cognitive and social development and the importance of understanding how children interpret their social cues. Therefore, it is important to investigate how children interpret the social signals of these agents, not only to assess their effectiveness but also to guide their design.

This line of research aligns with the growing evidence that specific features of humanoid and socially expressive robots, such as gaze, can enhance engagement as they are perceived to be more intentional, promoting social interactions (Wiese et al., 2017). Okumura et al. (2013) used Robovie robot and examined gaze effects on object preferences between human and robot agents. They revealed that 12-month-old infants preferred the objects gazed at by the human agent, while the robot gaze did not affect infants' object preferences. This results suggests that the gaze effects on the target object may depend on the mental states inferred from the agent's gaze (Manzi et al., 2023; Becchio et al., 2008). It is worthnothing that the lack of significant gaze effects from the robot agent could be partly explained by younger children's limited attribution of mental states to robots.

However, several studies showed that children from the age of five attribute some mental states to robots, although these attributions are generally less frequent and complex compared to humans (Di Dio et al., 2020; Manzi, Peretti, et al., 2020). For instance, children's prosocial behavior toward robots increases when robots are framed as capable of affective states, highlighting how mental state attributions can modulate children's responses to robotic agents (Nijssen et al., 2021). While children exhibit early forms of mental state attribution, it remains unclear how their developing Theory of Mind (ToM) abilities contribute to modulating their interpretation of gaze as an indicator of robot preference.

To date, no study has directly examined the relation between mental state attribution and the gaze effects on preferences in children. Furthemore, the gaze effects on preferences have been investigated in infants and adults, but not in children around the time of acquiring ToM. The abilities of mental state attributions have been measured in ToM tasks, and it has been suggested that children became acquiring ToM abilities between the ages of 3-5 years (Gopnik and Slaughter, 1991; Wellman et al., 2001). The ToM task in infants is limited to the implicit false-belief task, which can measure only infants' attributions of belief to others (Onishi and Baillargeon, 2005; Southgate et al., 2007). Using gaze information for own evaluation may involve inferring other's mental states rather than beliefs. By measuring individual differences in child ToM abilities using indexes other than the false-belief task, it is possible to investigate which aspects of mental state attributions are related to the gaze effects on preferences.

Moreover, using a robot agent could provide insight into the importance of mental state attributions on the gaze effects on preferences. The previous study on adults used an arrow cue to control the effects of attention, however, arrow cueing is qualitatively different from gaze cues (Bayliss et al., 2007). A recent meta-analysis has shown that the gaze cueing effect on attention is robust and similar across schematic, computer-generated, and real faces (McKay et al., 2021). On the other hand, arrow cues have weaker effects on attention compared to gaze cues (Friesen et al., 2004). Regarding to the results of the meta-analysis (McKay et al., 2021), robots with eye-like features are expected to orient human visual attention in the direction of their gaze (or facial movements). Thus, it is considered that a robot agent with eyes can lead attention as similar to a human agent compared to arrows. However, it remains unclear how robot gaze influences cognitive processing, such as preference toward the gaze target, beyond merely facilitating attention orientation. Another benefit of using the robot agent is that children do not spontaneously attribute mental states during observing gaze cueing situations (Manzi, Ishikawa, et al., 2020), therefore, by using robots in stead of arrows, it is possible to investigate whether the mental state attribution is necessary for the effects of gaze on object preference. Mental attributions to the Robovie are decreased compared to the human agent in 5-years-old children, but they are not entirely absent (Manzi, Peretti, et al., 2020). Social robots, like Robovie, can indeed elicit some level of mental state attribution in children, although these attributions are generally less robust and varied than those directed toward humans. By comparing gaze effects on object preferences between the robot and human agent and exploring correlation with the abilities of ToM respectively, the mechanisms based on mental state attributions can be examined.

To measure the gaze effects on object preference, the current study is based on the previous study by Okumura et al. (2013) examining how humanoid robot (i.e., Robovie) and human gaze affects infants' preference for objects. In Okumura's experiment, two objects were presented simultaneously, and either a human or the robot gazer directed their gaze toward one of the objects. The Robovie robot used in this study has a facial structure where head orientation and gaze direction are always aligned. Thus, the previous study examined whether head-turning consistent with gaze cueing increases preference for the object being looked at. Following prior research, we used similar gaze cueing scenarios with this same type of robot, focusing on preschool children as participants. In addition, our method for assessing preference attribution was adapted from Einav and Hood (Einav and Hood, 2006), who showed that 4- and 5-year-old children are able to infer an agent's preferred object based on gaze direction. This paradigm has been recently extended to adult participants (Manzi et al., 2025), who were able to correctly identify the preferred object of both human and robot agents. However, while adults could infer the agent's preference, the gaze cue did not influence their own explicitly stated preferences. The present study uses this paradigm to explore whether this pattern also characterizes early childhood.

The present study aims to assess the differential impact of human and robotic gaze on preschoolers' formation of preferences and their ability to attribute preferences to the gazing agent. Additionally, it explores how the attribution of mental states to agents and ToM abilities influence these processes. We hypothesize that preschool children will attribute preferences based on gaze cues more significantly when these cues are from human agents compared to robotic agents. This difference is expected due to the perceived capacity for genuine mental states in humans, which we posit is critical for influencing both the attribution and formation of preferences. Additionally, we anticipate that these gaze-induced preference attributions in children will correlate with their developmental progress in ToM abilities, particularly in discerning the mental states of others.

2. Methods

2.1. Participants

A total of 58 Italian children ($Mean_{age}$= 60.9 months, SD = 8.82 months; F = 29) were involved in the study. The sample estimate was conducted with G*Power to support the main analysis, i.e., a generalized repeated measures linear model (GLM). A sample estimate of 54 participants was established based on effect size f = 0.25, a probability of error = 0.05, a power estimate = 0.95. To account for potential no-shows or incomplete participation, the initial sample size included an additional 10% of participants, resulting in 4 extra children being tested and their data included in the final analysis. The parents were provided with a comprehensive description of the experimental process, tasks, and materials used during the research and provided written informed consent. The study was approved by the Research Ethics Review Board of the Department of Psychology, Kyoto University in Kyoto, Japan, with approval for this research (ethical proof number: 28-P12) and in accordance with the Declaration of Helsinki.

2.2. Robot

In the present study, we define a robot as an humanoid robot with anthropomorphic features necessary for conveying gaze as a social cue (e.g., head, eyes). This excludes non-anthropomorphic systems such as Roomba or Amazon's Alexa. Our definition follows prior work on humanoid robotics (Bartneck & Forlizzi, 2004; Ishiguro et al., 2001). In this sense, we used the Robovie robot (developed by Hiroshi Ishiguro Laboratories, ATR), an anthropomorphic humanoid robot designed with a mix of anthropomorphic and mechanical features. Robovie has two eyes and a microphone that resembles a mouth, enhancing its human-like qualities, but lacks legs (using two drive wheels for movement) and hands (with arms ending without hands). Its design falls within the mid-range of the mechanical-human continuum (see the ABOT database; Phillips et al., 2018), achieving a balance that allows it to resemble a human form while remaining distinctly non-human. This combination of features makes Robovie particularly suitable for studying social cues like gaze, as it presents children with a partially human-like agent to which they may attribute mental states, without fully realistic characteristics. Moreover, Robovie has been extensively used across different age groups, including children, to analyze gaze effects and mental state attribution in human-robot interactions (Okumura et al., 2013; Manzi, Peretti, et al., 2020).

2.3. Procedure

Each child was assessed individually in a quiet room within their kindergarten, which was deemed suitable for testing. The assessment session lasted approximately 20 min and was conducted by a researcher during school hours. The experimental protocol included: a gaze preference task, a first-order false belief ToM task, and a mental state attribution questionnaire for both the human and the robot. The sequence of these tasks was randomized for each child. For the execution of the gaze preference task, a Lenovo Yoga 530 touchscreen with a resolution of 1366x768 was utilized. This setup allowed the children to interact directly with the interface during the task, enabling them to intuitively select their responses to the stimuli presented. Further details about the instruments and methodologies employed in the assessment are provided in the subsequent paragraphs.

2.4. Design and gaze preference task

The study adopted a 2x2 repeated measures design, with two within-subject factors: *Agent* (Human, Robot) and *Perspective* (Gazer Preference, Participant Preference). The dependent variables were the attribution of gazer preference and individual preference, i.e., in correctly identifying the preferred object looked at by the gazer. Participants viewed six brief videos, each lasting 10 s, in which a human or a robot looked towards one of two objects. The face of the agent (human or robot) was displayed at the top of the screen, and two objects (the target and distractor) were shown at the bottom on a black background (see Fig. 1; for all targets and distractors see Appendix 1). The object pairings adopted in the present study were derived from a previous experiment conducted with adult participants (Manzi et al., 2025), in which a preliminary validation phase was used to select pairs with comparable preference levels. To ensure that these pairings were also neutral for preschool children, a pre-experimental preference check was conducted. At the beginning of the session, all objects were presented individually in a randomized order, and children were asked to rate how much they liked each object using a 5-point Likert smiley-face scale (1 = not at all; 5 = very much). Paired-sample *t*-tests indicated that no significant preference differences ($p > .05$) emerged within any object pair. Descriptive statistics and inferential results of this preference check are reported in Appendix 1 (Table A1). Across the experimental trials, the order of presentation of the six videos was fully randomized for each participant. Within each trial, the left/right position of the target and distractor objects as well as the direction of gaze were randomized. The agent initially made eye contact with the participant for 4 s before shifting gaze toward the target object and maintaining gaze for the remaining 6 s.

Before the start of the experimental trials, children underwent a familiarization phase. During this phase, children were randomly exposed to two trials of example with a pair of objects that were not used in the experimental trials. This phase was designed to familiarise participants with the agents and the visual format of the trials. Additionally, children were provided with the opportunity to independently interact with the Lenovo Yoga 530 touchscreen. This hands-on experience was meant to help children become familiar with the technology and complete the experiment independently, without the experimenter's assistance.

At the end of each video, participants were asked two questions presented in a counterbalanced order across trials and participants: "*Which object does the robot/girl like*?" (Gazer Preference) and "*Which object do you like*?" (Participant Preference). The first question followed previous paradigms of Einav and Hood (2006) and was used to explore whether children interpret gaze direction as a cue to preference. Congruent choices scored 1, i.e., when the participant's response matched the object gazed at by the agent (i.e., the target); incongruent choices scored 0, i.e., when the participant chose the distractor. The score for the Human and Robot conditions was calculated by adding the congruent choices from the three trials. Each child's responses were scored separately for the human and robot conditions, with a maximum score of 3 for each condition. The higher the score on the gaze preference task, the greater the impact of gaze on preference attribution.

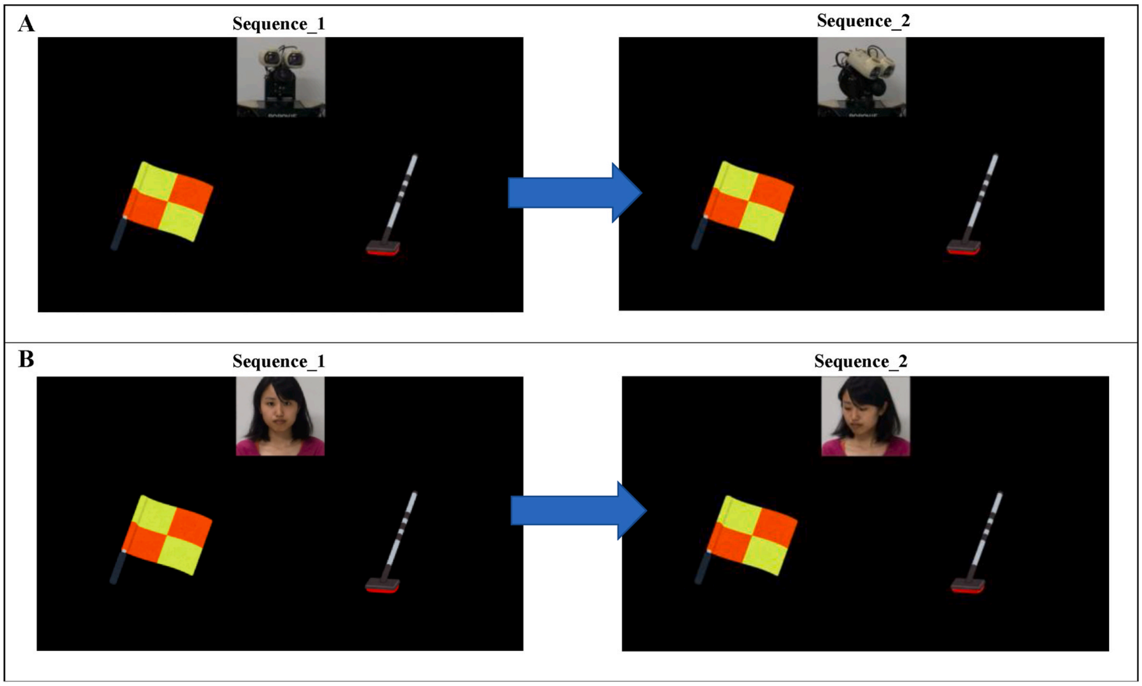


Fig. 1. A–B. The figure depicts two examples (A and B) of the video stimuli. In **A**, Sequence-1, the robot gazer made eye contact with the camera for 4 s, followed by looking at the target object for 6 s in Sequence-2. In **B**, Sequence-1, the human gazer made eye contact with the camera for 4 s, followed by looking at the target object for 6 s in Sequence-2.

2.5. *Theory of Mind – false belief*

To assess children's understanding of first-order ToM and their comprehension of false beliefs, the study employed the Unexpected Transfer task (Wimmer and Perner, 1983). During this task, a story was read to the child, featuring two characters (A and B), a ball, a box, and a basket, all situated in the same room. Character A places the ball in a box and then exits the room. Subsequently, Character B, in A's absence, moves the ball from the box to the basket. The child was then asked a series of questions by the experimenter regarding the sequence of events, including: (i) a false belief question (“Where does Character A think his ball is?”); (ii) a memory question (“Where did Character A originally place the ball before leaving the room?”); and (iii) a reality question (“Where is the ball actually located now?”). If the control questions (reality and memory) were answered accurately, the response to the false belief question was scored 1 if correct and 0 if incorrect.

2.6. *Attribution of Mental States Questionnaire (AMS-Q)*

The Attribution of Mental States Questionnaire (AMS-Q) is designed to assess the mental and perceptual states attributed by participants to specific characters, such as a human or a robot. In the present study, these attributions explicitly referred to the agents depicted in the experimental video stimuli. The questionnaire includes 23 items across various state categories including perceptual, emotional, desires and intentions, imagination, and epistemic states. Validated in its current form by Miraglia et al. (2023), the AMS-Q has been extensively utilized in research with both children (Di Dio et al., 2020; Manzi, Peretti, et al., 2020) and adults (Manzi et al., 2021). Participants respond to prompts such as, “In your opinion, can a human/robot think, decide, be happy, create, etc.?" Answers are binary, “Yes” or “No,” with “Yes” scored as 1 and “No” as 0. If a participant answers “Yes,” a follow-up question quantifies the response: “How much? A little bit or very much?”, allowing responses to be scored on a 3-point scale: No (0), Yes, a little bit (1), or Yes, very much (2). Reflecting its comprehensive validation in Italian, which identified three distinct factors, the AMS-Q employs a three-factor structure in this study: AMS-Neutral-Positive (AMS-NP; 13 items) evaluates the attribution of epistemic mental states, well-being states, and positive emotions (Human $\alpha = .9$; Robot $\alpha = .9$); AMS-Negative (AMS-N; 6 items) assesses states related to deception and negative emotions, also with an alpha of 0.9 for both agent types (Human $\alpha = .9$; Robot $\alpha = .9$); AMS-Sensory (AMS-S; 4 items) measures sensory state attributions (Human $\alpha = .9$; Robot $\alpha = .9$). The questionnaire was administered twice, once referring to the human agent and once to the robot agent. The order of administration (human first or robot first) was counterbalanced across participants. Each administration consisted of 23 items, identical for both agents, yielding a total of 46 responses per participant.

2.7. *Data analysis*

For the data analysis we used R statistical software (version 4.2.3). To examine the effects of gaze in the preference task, we carried out a Generalized Linear Model (GLM) with a binomial distribution, suited for handling the binary nature of responses (i.e., consistent or inconsistent choices across three fixed trials per condition; range: 0–3). This modeling choice aligns with prior recommendations for handling bounded count data (Agresti, 2013). Specifically, we employed Type II analysis of deviance to evaluate the effects in our experimental design involving interactions. This analytical approach was chosen for several key reasons. First, Type II analysis appropriately handles the hierarchical structure of effects by testing the interaction terms before examining main effects, which aligns with our experimental design where interaction effects are of primary interest. Second, this method provides a balanced treatment of main effects while controlling for other effects at the same level, reducing the risk of Type I errors compared to Type I analysis. Additionally, Type II analysis is particularly suitable for experimental designs as it evaluates each effect after controlling for all other effects of equal or lower order, allowing for a more conservative and reliable assessment of the experimental manipulations.

To examine the effects of ToM acquisition, we divided the children into two groups: children who passed the false-belief task ($n = 21$), and children who failed the false-belief task ($n = 37$). This model also integrated variables such as age, gender, agent type (human or robot), and perspective (gazer or participant). It also included an analysis of the interactions among these factors, allowing us to explore how these variables collectively affect children's preference attributions, particularly examining whether responses differ significantly between human and robotic gazers and whether children's perceptions align or conflict with the preferences of the agent.

Moreover, to identify potential relationships between the children's responses in the gaze preference task and the mental states they attribute to the agents, we performed correlation analyses between the AMS-Q scores for both human and robot, and the children's preference attributions to the gazer and their own preferences. Finally, a linear regression analysis was carried out to delve deeper into the factors that best predict children's preference attributions in the human condition.

3. Results

3.1. *Descriptive statistics*

Descriptive statistics for all measures are reported in Table 1, including gaze-preference scores for the human and robot conditions, False-Belief performance, and each AMS-Q subscale (AMS-NP, AMS-N, AMS-S) for both agents.

3.2. *Gaze preference task*

In the GLM, we modelled *Age, Gender, ToM, Agent* and *Perspective* and as fixed effects, and all combinations of factors as interaction terms (see Table 2). The results showed a significant interaction effect between *Agent* and *Perspective* ($\chi^2 = 11.643$, $p < .001$). Specifically, participants were more likely to attribute a preference to the gazer (rather than report their own preference) when the agent was human (estimate ± s.e. $= -0.814 \pm 0.220$, $Z = -3.669$, $p < .001$). Also, congruent attribution to the gazer's preference was higher for the human compared to the robot gazer (estimate ± s.e. $= -0.5801 \pm 0.2215$, $Z = -2.619$, $p = .043$). These findings are illustrated in Fig. 2. No other main effects or interaction terms reached significance. No significant relationship was found between gaze and participants' own preferences, regardless of whether the agent was human or robot.

Table 1
Descriptive statistics (means and standard deviations) for all measures.

		Mean	Std. Deviation
ToM		0.37	0.48
Gazer	*Human*	1.95	0.83
	Robot	1.53	1.04
Participant	*Human*	1.35	0.86
	Robot	1.70	0.82
AMS - Human	*NP*	1.43	0.50
	N	0.82	0.48
	S	1.53	0.55
AMS - Robot	*NP*	1.38	0.53
	N	0.68	0.44
	S	1.14	0.63

ToM = Theory of Mind (proportion correct); AMS = Attribution of Mental States; NP = Neutral Positive; N = Negative; S = Sensory.

Table 2

Statistical values of the deviance analysis (GLM type II tests).

Predictor	LR Chisq	p-value
Agent	0.100	0.752
Perspective	3.429	0.064
Gender	0.244	0.621
Age	0.943	0.332
ToM	0.102	0.749
Agent × Perspective	11.712	0.001***
Perspective × Gender	1.673	0.196
Agent × Gender	0.648	0.421
Agent × Age	0.199	0.655
Perspective × Age	1.709	0.191
Gender × Age	0.013	0.909
Gender × ToM	1.317	0.251
Age × ToM	0.264	0.607
Agent × ToM	0.003	0.958
Perspective × ToM	0.963	0.326
Gender × Age × ToM	0.653	0.419
Agent × Perspective × Gender	0.224	0.636
Agent × Perspective × Age	2.145	0.143
Agent × Gender × Age	0.041	0.839
Perspective × Gender × Age	0.244	0.621
Agent × Perspective × ToM	0.420	0.517
Agent × Perspective × Gender × Age	1.620	0.203

*Significance level codes: ***p < .001.

3.3. *Correlation with the attribution of mental states*

Pearson's correlation analysis between the AMS-Q scores for humans and robots, and the participants' own preferences and their attributions of the gazer's preferences revealed a significant positive correlations for the human, but not for the robot, $p > .05$, Specifically, the correlation between the attribution of gazer preference and the AMS-NP scores for humans was substantial ($r = 0.36$, $p = .008$). Indicating a moderate effect size. No significant correlations were observed for the robot condition ($p > .05$). These relationships are detailed in Table 3. Gender was coded as a binary variable (0 = male, 1 = female) and included in the correlation matrix for descriptive and exploratory purposes only. Additional non-parametric comparisons (i.e., Mann–Whitney *U* test) confirmed the absence of significant gender differences in preference scores or mental state attributions ($p > .005$). A paired-samples *t*-test on the total AMS-Q scores revealed that children attributed significantly more mental states to the human agent ($M = 3.8$, $SD = 1.11$) than to the robot agent ($M = 3.2$, $SD = 1.2$), $t\,(57) = 3.18$, $p = .002$, $d = 1.35$.

3.4. *Regression analyses*

To further examine the relationship between the attribution of preferences to the human gazer and the attribution of mental states to humans, we carried out a linear regression analysis, also considering demographic variables such as age and gender, as well as ToM and attribution of mental states (i.e., AMS-NP). This model was conducted only for the Human Gazer condition, as only this condition yielded a significant effect in the preference attribution task. Among the AMS-Q subscales, only the NP dimension was included as a predictor, based on its significant correlation with the dependent variable. This analysis aimed to assess whether the attribution of mental states to the human is associated with variation in children's preference attribution to the human gazer. The variables included are age, gender, ToM, and the AMS-NP (see Fig. 3). The final model identified AMS-NP as a significant predictor, explaining 36% of the variance in attribution of gazer preference (Adjusted R-squared = 0.12). The model was statistically significant, $F_{1,50} = 7.56, p = .008$. AMS-NP emerged as a robust predictor of preference attribution to the human gazer, $\beta = 0.4$, $t\,(56) = 2.8$, $p < 01$ (see Fig. 3). In contrast, age was not significantly associated with the children's response $\beta = -0.002$, $t\,(56) = -0.01$, $p > .05$, nor did gender, $\beta = 0.07$, $t\,(56) = 0.55$, $p > 05$, or ToM, $\beta = -0.004$, $t\,(56) = -0.02$, $p >$ .05.

Table 3

Pearson's r correlations between the participants' Age (months), Gender, ToM and mean preference scores in the Gazer and Participant perspective conditions for the Human and Robot.

		Human		Robot	
Correlation variables		Gazer	Participant	Gazer	Participant
Gender		0.065	0.117	−0.111	0.071
Age (months)		0.094	−0.183	−0.124	−0.190
ToM		0.002	0.091	−0.002	−0.139
AMS					
	NP	**.361**[a]	−0.164	0.017	0.047
	N	−0.224	−0.111	0.042	−0.010
	S	0.169	−0.091	0.218	0.114
N		58			

Pearson's r values are reported as effect sizes. ToM = Theory of Mind; AMS = Attribution of Mental States; NP = Neutral Positive; N = Negative; S = Sensory.

[a] Correlation is significant at the 0.01 level (2-tailed).

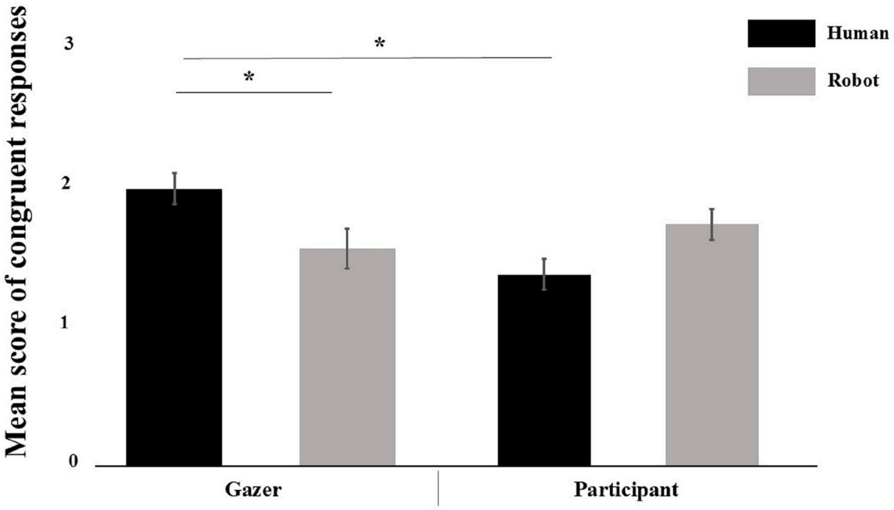


Fig. 2. Mean score of congruent responses in the Gazer and Participant preference tasks, separately for the Human (dark bars) and Robot (light bars) conditions.

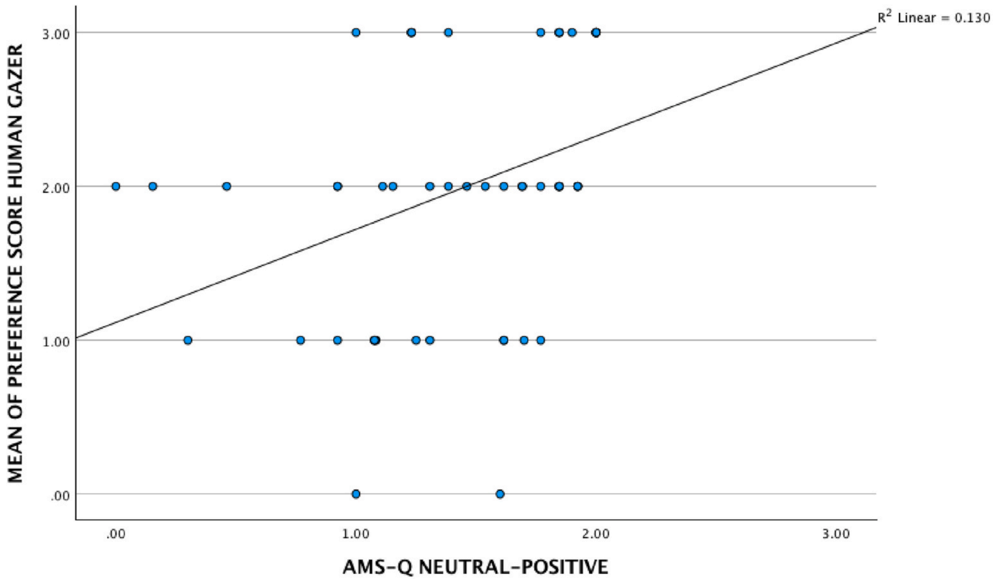


Fig. 3. Scatterplot showing a positive relationship between the attribution of preference to the human gazer (y-Axis) and the attribution of mental neutral-positive (x-Axis).

4. Discussion

The present study examined how preschool children interpret gaze cues from a human and a humanoid robot, focusing on two related but distinct outcomes: (1) whether children attribute preferences to the gazer based on gaze direction, and (2) whether those same gaze cues influence children's own preferences. Our findings revealed a significant difference in the attribution of preference between human and robot gazers, with human gaze exerting a substantial influence on the attributuion of preference to the gazer. Conversely, the influence of gaze on individual preference formation was not significant, regardless of the agent type. Furthermore, the regression analysis revealed a significant association between the attribution of mental states to the human and accuracy in the attribution of preferences to the human gazer. Finally, it is noteworthy that performance on the false-belief ToM task was not significantly associated with either the attribution of preferences to the gazer or the formation of individual preferences.

The marked difference in how children attribute preferences based on human compared to robot gaze reinforces existing evidence that early socio-cognitive processing is preferentially oriented toward human agents (Manzi, Ishikawa, et al., 2020; Okumura et al., 2013). In fact, preschool children seem to interpret human gaze not just as a simple directional signal, but as a communicative cue that conveys internal mental states, such as preferences or desires, in line with previous studies (Doherty et al., 2009; Einav and Hood, 2006). This suggests that children can explicitly associate the direction of human gaze with a preference (i.e., to desire something). In contrast, gaze from the humanoid robot did not significantly influence children's attribution of preferences. Importantly, this result should not be interpreted as indicating that robot gaze is generally uninformative. Rather, it suggests that gaze alone was not sufficient in the present task to support reliable preference attribution. This may reflect children's perception of robots as non-living entities. Supporting this, our AMS-Q data show significantly lower attribution of mental states to the humanoid robot than to the human, which aligns with findings typical for this age group (Manzi, Peretti, et al., 2020; Goldman and Poulin-Dubois, 2024; Di Dio et al., 2020). Although preschoolers may anthropomorphize robots to some extent, they may still have difficulty interpreting robot gaze as an expression of referential intentionality (i.e., a deliberate action aimed at directing attention toward an object to communicate a preference or desire) (Baron-Cohen et al., 1995; Frischen et al., 2007; Moll and Tomasello, 2007). This interpretation aligns with evidence indicating that, during the preschool years, the ability to infer preferences from gaze cues is still undergoing developmental consolidation (Doherty and Anderson, 1999; Doherty et al., 2009). As such, this skill may be reliably applied to human agents but not yet generalized to non-human entities such as robots. Supporting this, previous studies from infancy have shown a dissociation between gaze following and referential interpretation of robot gaze. Although infants visually follow the gaze of a humanoid robot, they do not reliably treat it as communicative or intentional (O'Connell et al., 2009). A similar pattern has been reported in children as young as two years old, who follow robot gaze more slowly than human gaze (Manzi, Ishikawa, et al., 2020), suggesting a developmental mismatch in responsiveness to social cues from human versus robotic agents. Importantly, this interpretation does not contradict previous evidence showing that preschool children can successfully learn from robots (e.g., Kory Westlund et al., 2017; Moriguchi et al., 2011). Rather, these findings suggest that such learning likely depends on the availability of additional communicative or contextual cues beyond gaze alone (e.g., verbal information or contingent interaction), which were minimized in the present paradigm. It is therefore likely that preschoolers continue to experience difficulties in ascribing desires to robot gaze while they are still refining this process in interactions with human agents. In a study using the same paradigm with adults, participants accurately inferred preferences from both a human and a humanoid robot gaze (Manzi et al., 2025). This seems to suggest a developmental change in the ability to interpret robot gaze as socially salient, probably supported by an increased experience that adults have of nonhuman agents, including robots. Another possible explanation is that adults may treat robots as representations or depictions of social agents rather than as genuinely animate entities, as suggested by recent theoretical accounts (Clark and Fischer, 2023). Identifying when and how this change occurs remains an important direction for future research.

The second main finding of the present study is that while children attribute preferences to the human based on gaze direction, their preferences are not influenced by the gaze observed by either the human or the robot. This difference supports recent theoretical perspectives that distinguish between attentional orientation and decision making (Orquin and Loose, 2013), suggesting that attention contributes to the decision making process but does not directly cause preference formation. Rather, decisions emerge from a complex interaction between attention, working memory, stimulus characteristics, and prior preferences. Although the attentional function of gaze is well documented in infants and adults (McKay et al., 2021; Moore, 2008), its influence on explicit decisions, especially in early childhood, has been less explored. Our results suggest that preschool children may decode gaze as a social cue to infer others' preferences, but this does not necessarily translate into changes in their own preferences. This pattern mirrors results obtained in adults using the same paradigm (Manzi et al., 2025), where participants also inferred the preferences of the gazer but were not influenced in their own choices. This may indicate that gaze, although sufficient for social inference (Doherty and Anderson, 1999; Doherty et al., 2009; Einav and Hood, 2006), is not a strong enough cue to alter explicit individual decisions. One possible explanation could be related to the novelty of the gazer agent. Previous research has shown that familiar individuals, such as parents or teachers, exert a greater influence on children's attention and learning, probably due to their established social relevance and trustworthiness (Barry-Anwar et al., 2017; Gredebäck et al., 2010; Hoehl et al., 2008). In our study, the human was unfamiliar to the children, which may have reduced the salience or persuasive power of the gaze. It is plausible that had a familiar adult been used, gaze cues might have had a greater effect on children's object choices, potentially altering the observed pattern of outcomes. This consideration becomes even more relevant when we consider the humanoid robot used in this study: not only was the robot unfamiliar, but it was also a novel and potentially ambiguous social agent, further limiting its influence on children's decision-making processes.

From a CCI perspective, these findings provide important constraints on how gaze should be conceptualized and implemented in child–robot interaction. While non-verbal behaviors such as gaze direction and minimal head movements are commonly treated as key design levers and can support learning and perception in child–robot interaction (Kory Westlund et al., 2017; Zaga et al., 2017), our results suggest that, for preschool children, gaze primarily supports social interpretation rather than direct behavioral influence. In other words, children may use gaze to infer what an agent prefers, but this does not necessarily translate into changes in their own decisions or preferences. This distinction has direct implications for the design of gaze-based interaction strategies: gaze alone may be insufficient to guide children's choices or learning outcomes, particularly when deployed by artificial agents, in the absence of additional communicative or contextual cues (e.g., verbal information or contingent interaction). Instead, gaze may need to be embedded within richer, contingent, and multimodal interaction contexts (e.g., verbal explanations, action consequences, or reciprocal feedback) to become an effective communicative signal in child–robot interaction, consistent with evidence that combining or varying non-verbal behaviors (e.g., gestures) can support robot engagement and learning over repeated interactions (De Haas et al., 2022), but also that deictic cues can backfire when poorly aligned with task demands (Aquilino et al., 2025; Yadollahi et al., 2018). Importantly, these findings highlight the need for developmentally informed design approaches that take into account both the socio-cognitive abilities of young users and their differential interpretation of human versus robotic agents (including humanoid robots such as the one used in the present study).

Finally, exploratory analysis of false belief understanding showed no significant association with either gazer attribution or their own preference. This result should be interpreted with caution, as only a relatively small proportion of children passed the False-Belief task, which may have limited variability in ToM performance and reduced statistical power to detect associations with gaze-based measures. Nevertheless, this pattern raises questions about the specific role of false belief reasoning in explicit gaze-mediated preference attribution. One possible explanation concerns the distinction between different forms of joint attention. While early joint attention abilities have been shown to predict later ToM development (Kristen et al., 2011), more recent work suggests that this relation is primarily driven by initiating joint attention, rather than responding to joint attention (Brandone and Stout, 2023). Given that the present task is based on gaze following, it is conceptually closer to responding to joint attention, for which links with ToM are less consistently observed, and this may help explain the absence of an association. Our data suggest that such attribution may not depend on the components involved in false belief reasoning, but rather on earlier-developing socio-cognitive mechanisms. This interpretation aligns with evidence indicating that gaze following and the attribution of desires or intentions emerge before the ability to represent others' false beliefs (Moll and Tomasello, 2007). In this sense, the explicit attribution of preferences may involve a less complex inferential process. Although evidence exists with respect to the association between gaze sensitivity in early life and the development of ToM measured with false beliefs (Brooks and Meltzoff, 2015), it is worth noting that a different dimension is measured in the present study: the explicit ability to attribute preferences to the other and to oneself. It will therefore be important in future studies to explore in more detail how these different components develop and interact.

In line with this perspective, our exploratory regression analyses showed that attribution of mental states to the human agent, as measured by the AMS-Q, significantly predicts gaze-based attribution of preferences. No similar association emerged for the robot, suggesting that the richness of mental reasoning used by children varies by agent. While exploratory, these results may suggest that explicit interpretation of social cues such as gaze may be supported by a broader range of socio-cognitive abilities than simply understanding false beliefs, including desires, emotions, and epistemic states. Further research is needed to examine how these dimensions jointly contribute to early social reasoning, and to identify developmental trajectories across different types of agents and tasks.

One limitation of this study is that it remains unclear whether children understand the gazer's object preference based on gaze direction alone. This study employed the same type of humanoid robot and similar gaze cueing scenarios (with two objects) as in previous study with infants (Manzi, Ishikawa, et al., 2020; Okumura et al., 2013). In such cases, understanding the gazer's preference requires a relative evaluation between two objects. However, it remains uncertain whether children can interpret preference from gaze direction when only a single object is presented. Differences in preference understanding based on the number of objects presented have not yet been explored, with previous studies focusing solely on differences in cued versus uncued effects regardless of the number of objects (Hoehl et al., 2008). Future research is needed to examine how the number of objects in gaze cueing scenarios influences preference interpretation from gaze direction in children. Another limitation is the difference between the facial features of the robot used in this study and those of a human. To examine the effect of gaze direction on preference understanding more directly, control conditions, such as "turning head with eyes closed," would be necessary. However, the humanoid robot used in this study (Robovie) and prior research (Manzi et al., 2023; Okumura et al., 2013) has a design in which head orientation and gaze direction are always aligned. This structural difference between human and robot faces may influence how gaze direction impacts children's cognitive processing. Future studies should explore how varying facial features in robots affect gaze interpretation, allowing for a comparison of the effects of anthropomorphism on gaze cues. In this respect, additional factors related to robot embodiment and task structure warrant further consideration. An additional aspect that may have influenced children's interpretation of robot gaze concerns the anthropomorphic properties of the robot itself. The humanoid robot used in the present study (Robovie) occupies a mid-range position on the anthropomorphism continuum, combining human-like facial features with clearly mechanical characteristics (Manzi, Peretti, et al., 2020). It is therefore possible that a more human-like robot, such as NAO, could elicit stronger attribution of intentionality and, consequently, different gaze-based preference patterns. This interpretation is consistent with recent evidence showing that preschool children tend to conceptualize robots primarily as artefacts rather than as animate beings (Di Dio et al., 2020; Goldman and Poulin-Dubois, 2024). For instance, Baumann et al. (2023) demonstrated that children of this age categorize robots as artefacts in an 'internal body part' task, which may constrain the attribution of complex mental states to robotic agents.

It is important to note that the present study relied on behavioral responses to infer gaze interpretation. Although children's choices indicate how gaze cues were interpreted at a decision level, future studies would benefit from integrating eye-tracking measures to directly assess attentional orienting and gaze following. Combining eye-tracking data with behavioral measures would allow a clearer dissociation between perceptual sensitivity to gaze direction and higher-level inferential processes related to preference attribution and mental state reasoning.

5. Concluding remarks and future directions

This study provides new insights into how preschool children process gaze cue from humans and robots, highlighting a clear asymmetry between the attribution of preferences to others and the formation of their own preferences. Our results confirm that preschoolers already interpret human gaze as a communicative cue that conveys mental states such as desires or preferences. However, this conclusion is not reached for robot, probably due to the lower intentionality attributed to it. Although robot gaze is perceptually similar, it does not seem to have any referential meaning for young children, suggesting an early developmental bias in the social cognition toward humans. Importantly, our results also indicate that gaze direction is sufficient to attribute preferences to others, but has no influence on children preferences. This difference, observed in both humans and robots, suggests that gaze cue alone may not be sufficient to modulate explicit decision-making in preschool children. This is consistent with recent theories that emphasize the multifactorial nature of gaze decision-making. Furthermore, the results of the exploratory regression suggest that a more comprehensive understanding of mental state, beyond false belief reasoning, plays a role in children attribution of preferences at least to humans. Although these results should be interpreted with caution, they underscore the importance of integrating multifactorial assessments of social-cognitive development, such as those captured by the AMS-Q.

From a Child–Computer Interaction perspective, these findings suggest that gaze should not be assumed to function as a stand-alone communicative or decision-shaping signal in child–robot interaction. Rather, gaze may primarily support social interpretation, and its effectiveness is likely constrained by children's developmental stage and their interpretation of the agent's intentionality. Consequently, gaze-based interaction strategies for young children may benefit from being embedded within richer, multimodal, and contingent interaction contexts, and from being informed by age-sensitive design principles.

Future studies should further investigate the developmental trajectory of these processes. In particular, longitudinal and age-comparative designs could help clarify when and how children start to interpret robot gaze as a social salient cue and. It would also be useful to replicate the current paradigm with familiar agents (e.g., parents, teachers) to test whether social familiarity increases the influence of gaze on preference formation. Furthermore, future studies should investigate how contextual cues and agent type modulate children's response to gaze, especially during interactions with artificial agents. Methodologically, future work should integrate eye-tracking measures to directly assess gaze following, allowing a clearer distinction between perceptual orienting to gaze and higher-level inferential processes involved in interpreting gaze as intentional. Finally, it will be important to discover which other cognitive and sociocognitive components may support these processes. For example, executive functions such as inhibitory control, cognitive flexibility, and working memory may interact with social cognitive mechanisms to shape children's ability to integrate gaze cues into decision making.

CRediT authorship contribution statement

Federico Manzi: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mitsuhiko Ishikawa:** Writing – original draft, Methodology, Formal analysis, Conceptualization. **Cinzia Di Dio:** Writing – original draft. **Shoji Itakura:** Writing – original draft. **Takayuki Kanda:** Writing – original draft. **Hiroshi Ishiguro:** Writing – original draft. **Davide Massaro:** Writing – original draft. **Antonella Marchetti:** Writing – original draft, Supervision.

Selection and participation

The research protocol was approved by the Research Ethics Review Board of the Department of Psychology, Kyoto University (Kyoto, Japan), in accordance with the Declaration of Helsinki (Approval No. 28-P12).

Participants were recruited from two preschools in the province of Milan, Italy. Following an initial approval of the research project by the school coordinators, the class teachers presented the study to the parents of all children in the relevant age group. The teachers provided parents with a comprehensive information sheet describing the experimental procedure, the tasks involved, and the materials that would be used.

Participation in the study was voluntary and contingent on receiving written informed consent from a parent or legal guardian. Only children for whom consent was provided were included in the research. The data collection sessions were conducted on the school premises during regular school hours, in a manner designed not to interfere with the standard educational activities. Children were informed that they could stop their participation at any time. All collected data were anonymized to ensure confidentiality and protect the identity of the participants.

Funding

The publication is supported by open access funding provided by Università Cattolica del Sacro Cuore within the CRUI-CARE Agreement. The research (F.M.) is supported by 'PON REACT EU DM 1062/21 57-I-999-1: Artificial agents, humanoid robots and human-robot interactions' funding of the Università Cattolica del Sacro Cuore of Milan. Furthermore, the research is supported by JST Moonshot R and D [grant number JPMJMS2011].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix 1

A. Stimuli used in the gaze preference task

This section of the Appendix presents all objects used as targets and distractors in the gaze-preference task. The object pairs were as follows.

Pair A) 1 - Flag, 2 – Broom
Pair B) 3 - Skittle, 4 – Tweezers;
Pair C) 5 - Tube, 6 - Party horn

B. Pre-experimental preference check of object stimuli

Prior to the gaze-preference task, children rated their liking for each object using a 5-point Likert smiley-face scale (1 = not at all; 5 = very much). Mean preference scores and results of paired-sample t-tests for each object pair are reported in Table A1. No significant differences emerged between the two items within any pair, indicating that the object pairings were neutral with respect to children's baseline preferences.

Table A1

Descriptive statistics and paired-sample t-tests for object preference ratings

	Object 1 (M ± SD)	Object 2 (M ± SD)	p (two-tailed)
Pair A: Flag (1) – Broom (2)	4.07 ± 1.27	3.65 ± 1.48	0.054
Pair B: Skittle (1) – Tweezers (2)	3.84 ± 1.37	3.54 ± 1.60	0.212
Pair C: Tube (1) – Party horn (2)	4.00 ± 1.20	3.63 ± 1.44	0.126

Data availability

Data will be made available on request.

References

Admoni, H., & Scassellati, B. (2017). Social eye gaze in human-robot interaction: A review. *J. Hum.-Robot Interact., 6*(1), 25–63. https://doi.org/10.5898/JHRI.6.1.Admoni

Agresti, A. (2013). *Categorical data analysis*. John Wiley & Sons. https://books.google.com/books?hl=it&lr=&id=6PHHE1Cr44AC&oi=fnd&pg=PR13&dq=Agresti,+A.+(2013).+Categorical+data+analysis.+John+Wiley+%26+Sons.&ots=KMipBjhP4I&sig=zcRUgaxyZdtKuOArV5-B4LBRxDQ.

Aquilino, L., Di Dio, C., Manzi, F., Massaro, D., Bisconti, P., & Marchetti, A. (2025, July). Decoding trust in artificial intelligence: A systematic review of quantitative measures and related variables. *Informatics, 12*(3), 70–110. https://doi.org/10.3390/informatics12030070

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology, 13*(4), 379–398. https://doi.org/10.1111/j.2044-835X.1995.tb00687.x

Barry-Anwar, R. A., Burris, J. L., Estes, K. G., & Rivera, S. M. (2017). Caregivers and strangers: The influence of familiarity on gaze following and learning. *Infant Behavior and Development, 46*, 46–58.

Bartneck, C., & Forlizzi, J. (2004). A design-centred framework for social human-robot interaction. *RO-MAN 2004. 13th IEEE international workshop on robot and human interactive communication (IEEE catalog No.04TH8759)*. https://doi.org/10.1109/ROMAN.2004.1374827

Baumann, A.-E., Goldman, E. J., Meltzer, A., & Poulin-Dubois, D. (2023). People do not always know best: Preschoolers' trust in social robots. *Journal of Cognition and Development, 24*(4), 535–562. https://doi.org/10.1080/15248372.2023.2178435

Bayliss, A., Frischen, A., Fenske, M., & Tipper, S. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition, 104*(3), 644–653. https://doi.org/10.1016/j.cognition.2006.07.012

Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review, 13*(6), 1061–1066. https://doi.org/10.3758/BF03213926

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Sciences, 12*(7), 254–258. https://doi.org/10.1016/j.tics.2008.04.005

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics, 3*(21). https://doi.org/10.1126/scirobotics.aat5954. eaat5954.

Brandone, A. C., & Stout, W. (2023). The origins of theory of mind in infant social cognition: Investigating longitudinal pathways from intention understanding and joint attention to preschool theory of mind. *Journal of Cognition and Development, 24*(3), 375–396. https://doi.org/10.1080/15248372.2022.2146117

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language, 35*(1), 207–220. https://doi.org/10.1017/S030500090700829X

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology, 130*, 67–78. https://doi.org/10.1016/j.jecp.2014.09.010

Çetinçelik, M., Rowland, C. F., & Snijders, T. M. (2021). Do the eyes have it? A systematic review on the role of eye gaze in infant language development. *Frontiers in Psychology, 11*, Article 589096. https://doi.org/10.3389/fpsyg.2020.589096

Clark, H. H., & Fischer, K. (2023). Social robots as depictions of social agents. *Behavioral and Brain Sciences, 46*, e21. https://doi.org/10.1017/S0140525X22000668

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Csibra, G., & Volein, Á. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology, 26*(1), 1–11. https://doi.org/10.1348/026151007X185987

De Haas, M., Vogt, P., Van Den Berghe, R., Leseman, P., Oudgenoeg-Paz, O., Willemsen, B., De Wit, J., & Krahmer, E. (2022). Engagement in longitudinal child-robot language learning interactions: Disentangling robot and task engagement. *International Journal of Child-Computer Interaction, 33*, Article 100501. https://doi.org/10.1016/j.ijcci.2022.100501

Di Dio, C., Manzi, F., Peretti, G., Cangelosi, A., Harris, P. L., Massaro, D., & Marchetti, A. (2020). Shall I trust you? From child–robot interaction to trusting relationships. *Frontiers in Psychology, 11*, 469. https://doi.org/10.3389/fpsyg.2020.00469

Doherty, M. J., & Anderson, J. R. (1999). A new look at gaze: Preschool children's understanding of eye-direction. *Cognitive Development, 14*(4), 549–571.

Doherty, M. J., Anderson, J. R., & Howieson, L. (2009). The rapid development of explicit gaze judgment ability at 3 years. *Journal of Experimental Child Psychology, 104*(3), 296–312.

Doherty, M. J., McIntyre, A. H., & Langton, S. R. H. (2015). Developmentally distinct gaze processing systems: Luminance versus geometric cues. *Cognition, 137*, 72–80. https://doi.org/10.1016/j.cognition.2015.01.001

Einav, S., & Hood, B. M. (2006). Children's use of the temporal dimension of gaze for inferring preference. *Developmental Psychology, 42*(1), 142–152. https://doi.org/10.1037/0012-1649.42.1.142

El-Muhammady, M. F., Zabidi, S. A. M., Yusof, H. M., Rashidan, M. A., Sidek, S. N., & Ghazali, A. S. (2023). In J. Jo, H.-L. Choi, M. Helbig, H. Oh, J. Hwangbo, C.-H. Lee, & B. Stantic (Eds.), *Initial response in HRI: A pilot study on autism spectrum disorder children interacting with a humanoid QTrobot* (Vol. 642, pp. 393–406). Springer International Publishing. https://doi.org/10.1007/978-3-031-26889-2_36. A c. Di), *Robot Intelligence Technology and Applications 7*.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews, 24*(6), 581–604. https://doi.org/10.1016/S0149-7634(00)00025-7

Esfandbod, A., Rokhi, Z., Meghdari, A. F., Taheri, A., Soleymani, Z., Alemi, M., & Karimi, M. (2023). Fast mapping in word-learning: A case study on the humanoid social robots' impacts on Children's performance. *International Journal of Child-Computer Interaction, 38*, Article 100614. https://doi.org/10.1016/j.ijcci.2023.100614

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy, 5*(1), 39–60. https://doi.org/10.1207/s15327078in0501_2

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance, 30*(2), 319–329. https://doi.org/10.1037/0096-1523.30.2.319

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*(4), 694–724. https://doi.org/10.1037/0033-2909.133.4.694

Ghiglino, D., Floris, F., De Tommaso, D., Kompatsiari, K., Chevalier, P., Priolo, T., & Wykowska, A. (2023). Artificial scaffolding: Augmenting social cognition by means of robot technology. *Autism Research, 16*(5), 997–1008. https://doi.org/10.1002/aur.2906

Giannakos, M., Papamitsiou, Z., Markopoulos, P., Read, J., & Hourcade, J. P. (2020). Mapping child–computer interaction research through co-word analysis. *International Journal of Child-Computer Interaction, 23–24*, Article 100165. https://doi.org/10.1016/j.ijcci.2020.100165

Goldman, E. J., & Poulin-Dubois, D. (2024). Children's anthropomorphism of inanimate agents. *WIREs Cognitive Science, 15*(4), Article e1676. https://doi.org/10.1002/wcs.1676

Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development, 62*(1), 98. https://doi.org/10.2307/1130707

Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: A longitudinal study of gaze following during interactions with mothers and strangers: The development of joint visual attention. *Developmental Science, 13*(6), 839–848. https://doi.org/10.1111/j.1467-7687.2009.00945.x

Hoehl, S., Wiese, L., & Striano, T. (2008). Young infants' neural processing of objects is affected by eye gaze direction and emotional expression. *PLoS One, 3*(6), Article e2389. https://doi.org/10.1371/journal.pone.0002389

Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T., & Nakatsu, R. (2001). Robovie: An interactive humanoid robot. *Industrial Robot: International Journal, 28*(6), 498–504. https://doi.org/10.1108/01439910110410051

Ishikawa, M., & Itakura, S. (2018). Observing others' gaze direction affects infants' preference for looking at gazing-or gazed-at faces. *Frontiers in Psychology, 9*, 1503. https://doi.org/10.3389/fpsyg.2018.01503

Ishikawa, M., Yoshimura, M., Sato, H., & Itakura, S. (2019). Effects of attentional behaviours on infant visual preferences and object choice. *Cognitive Processing, 20*(3), 317–324. https://doi.org/10.1007/s10339-019-00918-x

Kanda, T., Sato, R., Saiwaki, N., & Ishiguro, H. (2007). A two-month field trial in an elementary school for long-term human–robot interaction. *IEEE Transactions on Robotics, 23*(5), 962–971.

Konijn, E. A., Jansen, B., Mondaca Bustos, V., Hobbelink, V. L. N. F., & Preciado Vanegas, D. (2022). Social robots for (Second) language learning in (Migrant) primary school children. *International Journal of Social Robotics, 14*(3), 827–843. https://doi.org/10.1007/s12369-021-00824-3

Kory Westlund, J. M., Jeong, S., Park, H. W., Ronfard, S., Adhikari, A., Harris, P. L., DeSteno, D., & Breazeal, C. L. (2017). Flat vs. expressive storytelling: Young children's learning and retention of a social robot's narrative. *Frontiers in Human Neuroscience, 11*, 295. https://doi.org/10.3389/fnhum.2017.00295

Kristen, S., Sodian, B., Thoermer, C., & Perst, H. (2011). Infants' joint attention skills predict toddlers' emerging mental state language. *Developmental Psychology, 47*(5), 1207–1219. https://doi.org/10.1037/a0024808

Manzi, F., Ishikawa, M., Cinzia Di Dio, C., Itakura, S., Kanda, T., Ishiguro, H., Massaro, D., & Marchetti, A. (2025). Theory of mind abilities predict robot's gaze effects on object preference. *IEEE Transactions on Affective Computing.* https://doi.org/10.1109/TAFFC.2025.3531945

Manzi, F., Ishikawa, M., Di Dio, C., Itakura, S., Kanda, T., Ishiguro, H., Massaro, D., & Marchetti, A. (2020). The understanding of congruent and incongruent referential gaze in 17-month-old infants: an eye-tracking study comparing human and robot. *Scientific Reports, 10*(1), Article 11918. https://doi.org/10.1038/s41598-020-69140-6

Manzi, F., Ishikawa, M., Di Dio, C., Itakura, S., Kanda, T., Ishiguro, H., Massaro, D., & Marchetti, A. (2023). Infants' prediction of humanoid robot's goal-directed action. *International Journal of Social Robotics, 15*(8), 1387–1397. https://doi.org/10.1007/s12369-022-00941-7

Manzi, F., Peretti, G., Di Dio, C., Cangelosi, A., Itakura, S., Kanda, T., Ishiguro, H., Massaro, D., & Marchetti, A. (2020). A robot is not worth another: Exploring children's mental state attribution to different humanoid robots. *Frontiers in Psychology, 11*, 2011. https://doi.org/10.3389/fpsyg.2020.02011

Manzi, F., Sorgente, A., Massaro, D., Villani, D., Di Lernia, D., Malighetti, C., Gaggioli, A., Rossignoli, D., Sandini, G., Sciutti, A., Rea, F., Maggioni, M. A., Marchetti, A., & Riva, G. (2021). Emerging adults' expectations about the next generation of robots: Exploring robotic needs through a latent profile analysis. *Cyberpsychology, Behavior, and Social Networking, 24*(5), 315–323. https://doi.org/10.1089/cyber.2020.0161

McKay, K. T., Grainger, S. A., Coundouris, S. P., Skorich, D. P., Phillips, L. H., & Henry, J. D. (2021). Visual attentional orienting by eye gaze: A meta-analytic review of the gaze-cueing effect. *Psychological Bulletin, 147*(12), 1269–1289. https://doi.org/10.1037/bul0000353

Meltzoff, A. N., Brooks, R., Shon, A. P., & Rao, R. P. N. (2010). "social" robots are psychological agents for infants: A test of gaze following. *Neural Networks, 23*(8–9), 966–972. https://doi.org/10.1016/j.neunet.2010.09.005

Miraglia, L., Peretti, G., Manzi, F., Di Dio, C., Massaro, D., & Marchetti, A. (2023). Development and validation of the Attribution of Mental States Questionnaire (AMS-Q): A reference tool for assessing anthropomorphism. *Frontiers in Psychology, 14*, Article 999921. https://doi.org/10.3389/fpsyg.2023.999921

Moll, H., & Tomasello, M. (2007). Cooperation and human cognition: The Vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society, B.* https://doi.org/10.1098/rstb.2006.2000

Moore, C. (2008). The development of gaze following. *Child Development Perspectives, 2*(2), 66–70. https://doi.org/10.1111/j.1750-8606.2008.00052.x

Moriguchi, Y., Kanda, T., Ishiguro, H., Shimada, Y., & Itakura, S. (2011). Can young children learn words from a robot? *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems, 12*(1), 107–118. https://doi.org/10.1075/is.12.1.04mor

Nijssen, S. R. R., Müller, B. C. N., Bosse, T., & Paulus, M. (2021). You, robot? The role of anthropomorphic emotion attributions in children's sharing with a robot. *International Journal of Child-Computer Interaction, 30*, Article 100319. https://doi.org/10.1016/j.ijcci.2021.100319

O'Connell, L., Poulin-Dubois, D., Demke, T., & Guay, A. (2009). Can infants use a Nonhuman agent's gaze direction to establish word–object relations? *Infancy, 14*(4), 414–438. https://doi.org/10.1080/15250000902994073

Okumura, Y., Kanakogi, Y., Kanda, T., Ishiguro, H., & Itakura, S. (2013). Infants understand the referential nature of human gaze but not robot gaze. *Journal of Experimental Child Psychology, 116*(1), 86–95. https://doi.org/10.1016/j.jecp.2013.02.007

Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old infants understand false beliefs? *Science, 308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144*(1), 190–206.

Parsons, J. P., Bedford, R., Jones, E. J. H., Charman, T., Johnson, M. H., & Gliga, T. (2019). Gaze following and attention to objects in infants at familial risk for ASD. *Frontiers in Psychology, 10*, 1799. https://doi.org/10.3389/fpsyg.2019.01799

Phillips, E., Zhao, X., Ullman, D., & Malle, B. F. (2018). What is human-like?: Decomposing robots' human-like appearance using the anthropomorphic roBOT (ABOT) database. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction* (pp. 105–113). https://doi.org/10.1145/3171221.3171268

Senju, A., Csibra, G., & Johnson, M. H. (2008). Understanding the referential nature of looking: Infants' preference for object-directed gaze. *Cognition, 108*(2), 303–319. https://doi.org/10.1016/j.cognition.2008.02.009

Shiomi, M., Kanda, T., Howley, I., Hayashi, K., & Hagita, N. (2015). Can a social robot stimulate science curiosity in classrooms? *International Journal of Social Robotics, 7*(5), 641–652. https://doi.org/10.1007/s12369-015-0303-1

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-Year-Olds. *Psychological Science, 18*(7), 587–592. https://doi.org/10.1111/j.1467-9280.2007.01944.x

Striano, T., Kopp, F., Grossmann, T., & Reid, V. M. (2006). Eye contact influences neural processing of emotional expressions in 4-month-old infants. *Social Cognitive and Affective Neuroscience, 1*(2), 87–94. https://doi.org/10.1093/scan/nsl008

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

Wiese, E., Metta, G., & Wykowska, A. (2017). Robots as intentional agents: Using neuroscientific methods to make robots appear more social. *Frontiers in Psychology, 8*, 1663. https://doi.org/10.3389/fpsyg.2017.01663

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Yadollahi, E., Johal, W., Paiva, A., & Dillenbourg, P. (2018). When deictic gestures in a robot can harm child-robot collaboration. In *Proceedings of the 17th ACM conference on interaction design and children* (pp. 195–206). https://doi.org/10.1145/3202185.3202743

Zaga, C., De Vries, R. A. J., Li, J., Truong, K. P., & Evers, V. (2017). A simple nod of the head: The effect of minimal robot movements on children's perception of a low-anthropomorphic robot. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 336–341). https://doi.org/10.1145/3025453.3025995